Consortium

CGIAR

CONCEPT NOTE:
Providing Implementation Support for the CGIAR Open Access & Data Management Roadmap

I. EXECUTIVE SUMMARY

CGIAR produces a veritable goldmine of critical research data, information and knowledge. Our desire is to make that data, information, and knowledge more accessible to the world, to increase the pace of innovation and impact of the research enterprise. CGIAR has committed to making its research outputs as widely accessible as possible and in a timelier manner. In 2012, the CGIAR Consortium Board of Trustees approved "The CGIAR Principles on the Management of Intellectual Assets" requiring that "the Consortium . . . promptly and broadly disseminate their research results", and the Consortium Open Access Policy was approved by the CGIAR Consortium Board on October 2, 2013. The time is right to now support CRPs and their implementing Centers to implement this Policy, to collaborate on this initiative with other CGIAR entities including ISPC and IEA, and for CGIAR to take a leadership role in the overall Open Data for Agriculture movement.

This project's goal is to have 80% of all Open Access-eligible CGIAR Information Products accessible from a single portal within 3 years, and 100% available within 5 years. The roadmap will depend on a clear Policy and detailed Implementation Guidelines; close collaboration with partners on standards and data; widespread culture change through proactive human capacity development; robust technology platforms; and capacity-building and resource support for CRPs and the Centers and partners that implement them. A number of donors have expressed their support for this endeavor, and their assistance will be critical to success. As discussed later in this Concept Note, the project is estimated to require approximately $17M over the course of 5 years.

II. BACKGROUND/PROBLEM/OPPORTUNITY

Significant and pervasive limitations in the accessibility, timeliness, documentation and interoperability of CGIAR's data and knowledge assets are reducing research efficiency, raising research costs and generally stifling innovation both within and beyond CGIAR. Each year CGIAR Research Programs and Centers spend hundreds of millions of dollars generating scientific data, innovations, and knowledge products. Absent past leadership or effective policies with respect to the curation and sharing of these assets (data in particular), they have been significantly undervalued and underutilized. This represents a high opportunity cost in terms of the learning and innovation foregone, as well as a failure to fully render the public goods intended by the funders of the CGIAR system. This project aims to address this challenge head-on and to help transform CGIAR into an Open Access organization.

Recognizing its poor track record and acknowledging both changing scientific norms and the growing pressure from donors to deliver the full public-good benefits of their investments, the CGIAR Consortium has committed to making its research outputs as widely accessible as possible. Additionally, this project will allow CGIAR to take a leadership role in the sector-wide effort, highlighted by the G8 Summit on Open Data for Agriculture in April 2013, to play a pivotal and catalytic role in developing an entire "open" ecosystem for agricultural research and data.

Consortium

CGIAR

III. THE SOLUTION

The solution is one that should have widespread impact on the agricultural research for development sector, namely turning CGIAR into a truly "Open Access" entity, and adopting and supporting open standards. Progress has already been made. In addition to the April 2012 *CGIAR Principles on the Management of Intellectual Assets*, the CGIAR Consortium has developed an Open Access and Data Management Policy that has undergone widespread review with Centers, CRPs, Fund Council Members, the ISPC, and partners; the Policy was approved by the CGIAR Consortium Board on October 2, 2013.

But Principles and Policies are not enough. Now is the time for action. With appropriate resources to match the commitment of CGIAR Consortium members, we aim to make all of our information products openly accessible, universally harvestable, and completely interoperable within 5 years through a single Web-based portal that draws from any number of repositories. Through this project the CGIAR Consortium – Research Programs, Centers, and the Consortium Office – will implement the Open Access policy. The project will facilitate cross-CGIAR/cross-partner dialogue and agreement on the adoption of coherent, Open Access standards and best practices; facilitate improved approaches for incentivizing scientists to better document, curate and share their research outputs; develop robust platforms; and build capacity in CRPs and Centers.

IV. PROPOSED PROJECT DETAILS/ACTIVITIES

The CGIAR Open Access and Data Management Roadmap focuses on 7 key elements. Monitoring and Evaluation (M&E) metrics for each are included in section V, below.

1. ***Drawing the Map*: Guidelines and Standards.** The foundation for implementation begins with collaboratively developing Implementation Guidelines and identifying – and adding to, as necessary – the appropriate sector-wide standards to use that will best enable interoperability. This will involve establishing the core team from across the Consortium, working closely with them, and collaborating with other key agricultural data partners.

2. ***Making the Case:* Awareness and Culture Change.** One challenge to moving towards Open Access is the need to create an Open Access culture, one that incentivizes researchers to make their data and information openly accessible. Academic institutions and research organizations around the world are struggling with this challenge. A strong campaign that focuses on the "what" and "why" of Open Access will be critical, as will working with HR leaders to change the way that researchers are incentivized; these may include updating contracts, public recognition or an annual award for most viewed/downloaded open publications, tracking and reporting of downloaded data.

3. ***Building Capacity:* Center/CRP Human and IT capacity:** The OA Implementation Guidelines, already in draft, posted on our website for a first round of consultation, will require each CRP and Center to have adequate staffing to comply with the Policy, including dedicated data managers. They will also require building and managing compliant repositories that will allow for open harvesting of the deposited information products. Staffing up and building this IT capacity will require resources not currently budgeted for, as well as time to build that capacity.

Consortium

CGIAR

4. ***Building the Delivery Mechanism:*** **IT Platforms.** Once open standards have been agreed do, and the CRPs and Centers will have the capacity to deposit information products into their repositories, there needs to be a single interface into which such data, as well as those of partners using the same standards, can be harvested. We envision a portal at open.cgiar.org which will serve that purpose. It will be a meta repository for publications, associated data, as well as spatial, crop, weather, and landscape data. We will work with data management partners and contractors to build this portal.

5. ***Pulling it all together:*** **Implementation.** As Table 1 below indicates, we envision an implementation ramp-up that starts with 2 Centers and 3 CRPs in the first year, 10 Centers and 16 CRPs by the end of the third year, and all CRPs and Centers within 5 years. By the end of the third year, 80% of all eligible information products (IPs) will be openly accessible, with 100% of eligible IPs available within 5 years. Implementation will include working closely with partners to make sure our data and theirs is interoperable, and available on open.cgiar.org.

6. ***Leading the Way*****: Sector Leadership.** The agricultural research for development sector has galvanized around the concept of Open Data, and CGIAR now has the opportunity to step into a leadership role. Working with partners and across CGIAR including the ISPC and IEA, and across the sector, we can demonstrate the value of our research by making it available and keeping the momentum going. As part of this, we propose to be a lead partner for an annual Open Data for Agriculture conference which would pick up the mantle of the 2013 G8 Summit.

7. ***Testing the Theory:*** **Impact Assessment.** Underpinning all of this effort is the theory that more open and accessible data and information will lead to better decision-making, more effective investments, and improved eventual impact. We would like to conduct a serious impact assessment study in parallel with this project, in partnership with CGIAR's ISPC and IEA.

V. OUTCOMES/OBJECTIVES

Our ultimate goal is to have 80% of all eligible CGIAR Information Products in a format that can be accessible from a single portal within 3 years, and 100% available within 5 years. Table 1 below lays out a timeline for, and metrics to measure progress towards, the following 5-year objectives:

1. Guidelines established, adopted, and periodically reviewed. Sector-wide standards adopted, reviewed, adjusted as necessary.
2. Open Access becomes business as usual for CGIAR; CGIAR is known as a leader in transparent, openly accessible Agricultural Research for Development (AR4D)
3. 16 CRPs and 15 Centers have the human and technological capacity to comply with the Open Access Policy and Guidelines.
4. Open.cgiar.org is developed, and becomes the most visited, trusted, and interoperable Open Data for Ag resource on the Web.
5. 16 CRPs and 15 Centers are fully compliant with OA Guidelines. 80% of all eligible Information Products are openly accessible within 3 years, 100% within 5 years.
6. AR4D data widely interoperable across the sector; decision-making based on better, more available data is improved.
7. The impact of the connection between open data and development outcomes is assessed.

Consortium

CGIAR

TABLE 1: High Level Activities, Outputs, and Outcomes

Activities	Outcomes	Outputs/Targets					Outcomes *(Within 5 Years . . .)*
		Year 1	**Year 2**	**Year 3**	**Year 4**	**Year 5**	
Guidelines and Standards	Implementation Guidelines developed	Established	Reviewed	**Updated**	Reviewed	Updated	Guidelines established, adopted, and periodically reviewed. Sector-wide standards adopted, reviewed, adjusted as necessary.
	Standards adopted	Adopted	Reviewed	**Updated**	Reviewed	Updated	
Awareness and Culture Change	# Centers/CRPs receiving Capacity Building training	3 Centers 6 CRPs	9 Centers 16 CRPs	**15 Centers** **16 CRPs**			Open Access is business as usual; CGIAR is known as a leader in transparent, openly accessible AR4D
	# Centers, CRPs with OA HR incentives/metrics	2 Centers 3 CRPs	9 Centers 16 CRPs	**15 Centers** **16 CRPs**			
Center/ CRP Human/ IT Capacity	# Centers, CRPs with OA-dedicated staff	2 Centers 3 CRPs	6 Centers 6 CRPs	**10 Centers** **16 CRPs**	12 Centers 16 CRPs	15 Centers 16 CRPs	15 Centers and 16 CRPs have full OA compliance capacity.
	# Centers, CRPs with compliant repositories	2 Centers 3 CRPs	6 Centers 6 CRPs	**10 Centers** **16 CRPs**	12 Centers 16 CRPs	15 Centers 16 CRPs	
Platforms	Open.cgiar.org portal developed	Develop, pilot	Scale	**Review**	Upgrade	Review	Open.cgiar.org is most visited, trusted, and interoperable Open Data for Ag resource on the Web.
	% visits increase to open.cgiar.org	Baseline	50%	**50%**	50%	50%	
	% downloads increase from open.cgiar.org	Baseline	50%	**50%**	50%	50%	
	# partners harvesting from open.cgiar.org	0	5	**15**	25	50	
Implementation	# Centers, CRPs implementing Guidelines	2 Centers 3 CRPs	6 Centers 6 CRPs	**10 Centers** **16 CRPs**	12 Centers 16 CRPs	15 Centers 16 CRPs	15 Centers, 16 CRPs are fully compliant with OA Guidelines. 100% of all eligible Information Products are openly accessible.
	% of eligible IPs openly available.	25%	50%	**80%**	90%	100%	
Sector Leadership	# Open Data for Ag Conferences held	1	1	**1**	1	1	AR4D data widely interoperable; decision-making improved.
Impact Assessment	IA Study Conducted	Design	Track	**Track**	Track	Assess	Impact of connection between open data and impact assessed

Consortium

CGIAR

VI. CONCLUSION

The CGIAR Commitment towards Open Access; sector-wide momentum towards Open Data; a shifting landscape in research funding; and current information and communication technology have all converged to provide a moment in time that needs to be capitalized. Now is the time to invest in opening up CGIAR's research assets.

Open Access efforts are already happening across the Consortium, in publishing and in data. Good examples abound; what needs to happen now is institutionalization and standardization. But it will come with costs. To give a point of reference, in the past three years, an estimated $2M has been spent by Centers and CRPs to make a portion of their publications openly accessible -- just with one publisher[1]. Outside CGIAR, many organizations are investing heavily in Open Access and Open Data. Another example: The European Commission has committed to continue to fund projects related to open access; the Commission plans to spend €45 million on data infrastructures and research on digital preservation in 2012-2013 alone[2].

But while the will is there, implementation of the Open Access and Open Data roadmap will be a serious endeavor. In fact, while other organizations have produced Open Access policies and Implementation Guidelines, these are mostly focused on Open Access publishing; what CGIAR is proposing to do is actually a groundbreaking effort: producing a policy and implementation roadmap that addresses not only Open Access publishing, but openly accessible data. What's more, the effort will require a complete culture shift within and across CGIAR. Taking this leadership role will position CGIAR to help guide the global effort towards Open Data for Agriculture, to improve decision-making, research and, ultimately, impact. To help bolster this effort, we will conduct an accompanying impact assessment to understand the degree to which increased access to information leads to greater impact on the ground.

This project will devote more than 70% of its budget to CRPs and their implementing Centers to build capacity and implement the policy; nearly 18% to partners for capacity building, collaboration, and harmonization; and approximately 11% to the CGIAR Consortium Office to coordinate and lead. With appropriate focus and drive, CGIAR can be a truly "Open Access" organization within five years, the bulk of which can be accomplished within three. This will require bold action, a clear roadmap, and measurable metrics (see Appendix A). Key partners will be central to the process, including donor agencies such as USAID, DFID, GIZ, the Bill and Melinda Gates Foundation; key global data partners such as the World Bank and FAO; national and regional agricultural organizations such as USDA, FARA, Embrapa, and CAAS; and universities and other research hubs.

Collaborators from across the CGIAR Consortium stand ready to provide a more detailed proposal upon request.

[1] This data is still being collected; it will be compiled and presented in the full proposal, if requested.
[2] http://europa.eu/rapid/press-release_IP-12-790_en.htm

Consortium

CGIAR

VII. INDICATIVE BUDGET

The estimated budget for this 5-year initiative is shown in Table 2 Below. Of the estimated $16.9M required over 5 years, 71% would go to CRPs and their implementing Centers to build their capacity and to carry out the implementation; nearly 18% will go to partners and contractors; and 11% will go to the Consortium Office, whose job is will be to coordinate guide the process. The majority of funding comes in the first two years, after which point there is a significant decline. At the same time, we plan for real scale up of compliance by CRPs and implementing Centers, and corresponding information product availability (as shown in Table 1 above).

The project's funding needs are divided into five main categories (see table 2 below):

1. **Building the Foundation:** includes establishing cornerstones towards implementation, including establishing the core team; developing and adopting guidelines and standards; and building awareness and culture change. We envision this capacity to be integrated into Consortium components after 3 years, with no further requirement for external funding.
2. **CRP Capacity Building:** includes sub-grants to CRPs for hiring key staff and establishing compliant repositories. Support will only being required for two years, because CRP 2nd call Guidance will include how to budget/staff appropriately.
3. **CRP Implementation Support:** includes supporting the implementation and management of open access of all Centers supporting CRPs, including staff costs; building and managing compliant online repositories; publishing and data preparation fees; and internal practices, incentives, and capacity building.
4. **"Innovation Fund":** to support partners, particularly in developing countries, to build their capacity and support creative and complementary Open Access delivery mechanisms.
5. **Sector Leadership:** includes support for the annual Open Data for Agriculture conferences and the Impact Assessment study.

TABLE 2: Budget

Table 2: Funding	Year 1	Year 2	Year 3	Year 4	Year 5	TOTAL	CRPs, Centers	Partners, contractors	Consortium Office
Building the Foundation	1,800,000	1,670,000	1,500,000			**4,970,000**	1,500,000	1,500,000	1,970,000
CRP capacity building	1,500,000	1,500,000				**3,000,000**	3,000,000		
Imple-mentation support	1,500,000	1,500,000	1,500,000	1,500,000	1,500,000	**7,500,000**	7,500,000		
Innovation Fund	250,000	250,000	250,000	250,000		**1,000,000**		1,000,000	
Sector Leadership	100,000	100,000	100,000	100,000	100,000	**500,000**		500,000	
Totals	**5,150,000**	**5,020,000**	**3,350,000**	**1,850,000**	**1,600,000**	**16,970,000**	**12,000,000**	**3,000,000**	**1,970,000**

Consortium

CGIAR

APPENDIX 1: Monitoring and Evaluation Metrics

The following is a set of draft M&E metrics that will be reviewed with CRPs and Centers to assess the degree to which they are complying with the Open Access and Data Management Policy and Implementation Guidelines.

1. Number of final Information Products produced during the calendar year.
    a. Journal articles
    b. Datasets associated with journal articles
    c. Non-journal article-associated datasets
    d. Reports
    e. Videos
    f. Photographs
    g. Films
    h. Software products
    i. Other information products
2. Number of final Information Products eligible under CGIAR Open Access Policy (total number, minus Information Products that are justifiably excluded/aggregated or protected by intellectual property contract language, or not covered under the SRF)
3. Percent of eligible Information Products that are Open Access (See Table 4 below):
    a. % that meet 3 Star rating (number of 3-start Information Products divided by Number of eligible Information Products under CGIAR Open Access Policy)
    b. % that meet 2 Star rating.
    c. % that meet 1 Star rating.
4. Number of downloads of Open information products.
5. Number of citations of Open information products (traditional print media and data).
6. Number of partners harvesting Open information products.

Table 3: Monitoring and Evaluation Targets over 5 Years

	Year 1	Year 2	Year 3	Year 4	Year 5
# of IPs produced	#	#	#	#	#
# of IPs produced that are OA eligible	65%	70%	75%	80%	90%
% of OA IPs produced that are available	25%	50%	80%	90%	100%
# downloads of Open IPs	Baseline	50% increase	50% increase	50% increase	50% increase
# citations of Open IPs	Baseline	50% increase	50% increase	50% increase	50% increase
Partners harvesting Open IPs	0	5	15	25	50

Table 4: The "CGIAR Open Access Star Rating System"

1 Star	2 Star	3 Star
The Information Product (IP) sits on a website that doesn't require credentials, but does not comply with any data management standards.	The IP complies with CGIAR OA Guidelines in that it is stored in an openly accessible repository, uses a set of data standards, but it does not yet utilize agreed-upon metadata elements and adopted Common vocabulary.	The IP in question complies fully with CGIAR's Open Access Implementation Guidelines, including the use of agreed-upon metadata elements and adopted Common Vocabulary, and thereby available on open.cgiar.org.